Carcicast: Developing a Carcinogenicity Testing Toolbox

Nicole C. Kleinstreuer, PhD
Deputy Director, NICEATM
PI, Computational Toxicology Group, DIR/BCBB
National Institute of Environmental Health Sciences

29th April, 2019

Converging on Cancer Workshop

- National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)
- Division of Intramural Research, Biostatistics and Computational Biology Branch

UNITED STATES ICCVAM *Advancing Alternatives to Animal Testing*

National Toxicology Program NTP

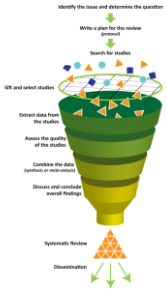
EPA/NCCT

Amgen

Disclaimer: Thoughts and examples provided herein are my own opinions. Mention of trade names or technologies does not constitute endorsement.

Constructing "CarciCast"*

- Expert-driven approach
  - Identify assays/biomarkers that map to key characteristics of carcinogens/hallmarks of cancer



- Semi-supervised systematic review
  - Broad keyword search for all relevant scientific literature, abstract screening and tagging

- Applying HTS data
  - Prioritize environmental chemicals based on bioactivity against targets that map to cancer hallmark pathways
  - Construct QSAR models for key characteristics

*Naming Credit: Keith Houck, EPA/NCCT

Expert-driven Approach

The Hallmarks of Cancer

- A. Self-sufficiency in growth signals
- B. Insensitivity to anti-growth signals
- C. Evading programmed cell death
- D. Limitless replicative potential
- E. Sustained angiogenesis
- F. Tissue invasion and metastasis

Original Hallmarks

- G. Deregulated metabolism
- H. Evading the immune system

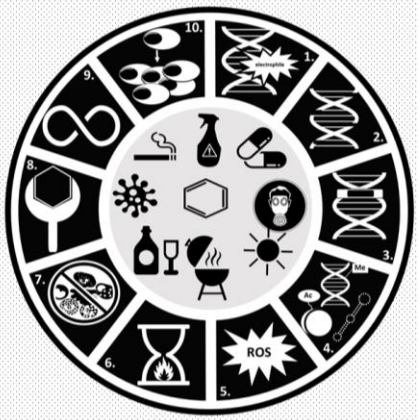
Emerging Hallmarks

- I. Genome instability
- J. Inflammation

Enabling Hallmarks

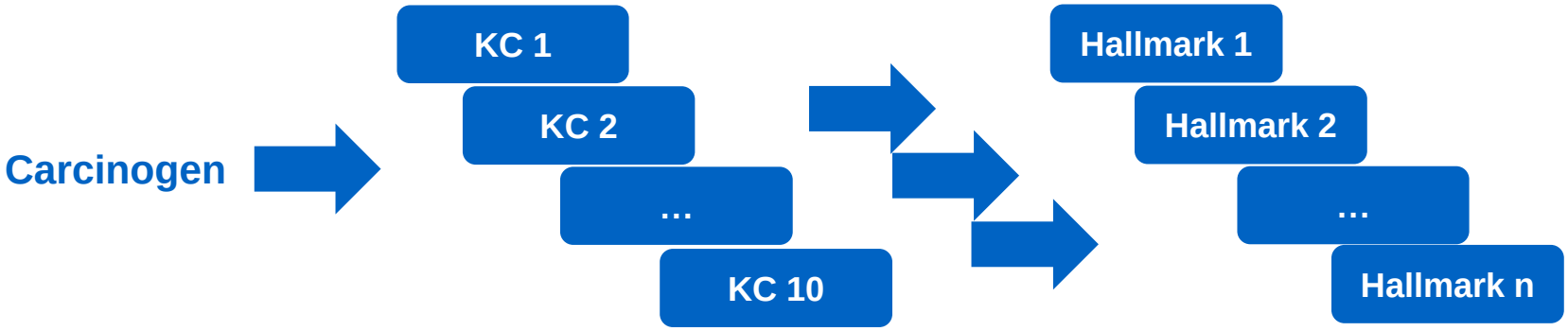
The Key Characteristics of Human Carcinogens

1. Is Electrophilic or Can Be Metabolically Activated to Electrophiles
2. Is Genotoxic
3. Alters DNA repair and Causes Genomic Instability
4. Induces Epigenetic Alterations
5. Induces Oxidative stress
6. Induces Chronic Inflammation
7. Is Immunosuppressive
8. Modulates Receptor-mediated effects
9. Causes Immortalization
10. Alters Cell Proliferation, Cell Death or Nutrient Supply



- Understanding the relationship between hallmarks (HM: *biology*) and key characteristics (KC: *chemistry*)
- Which are measurable, and in what platforms?
- KC: Chemical properties with associated targeted assays
- HM: Biological properties requiring integrated models

No Clear One-to-one Relationship

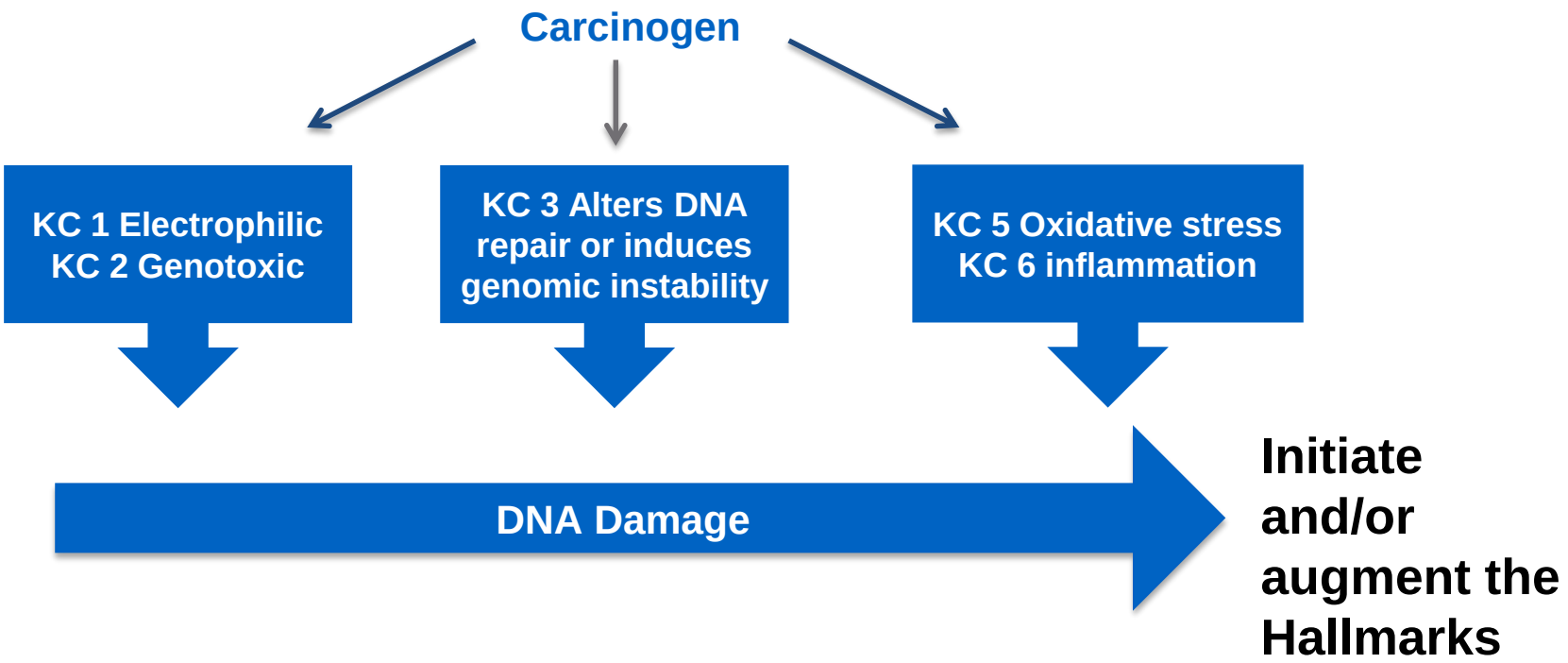


Carcinogens induce one or more KC's at one or more points in the process (i.e. initiation/promotion)

Tumors acquire one or more HMs at various points in the carcinogenic process

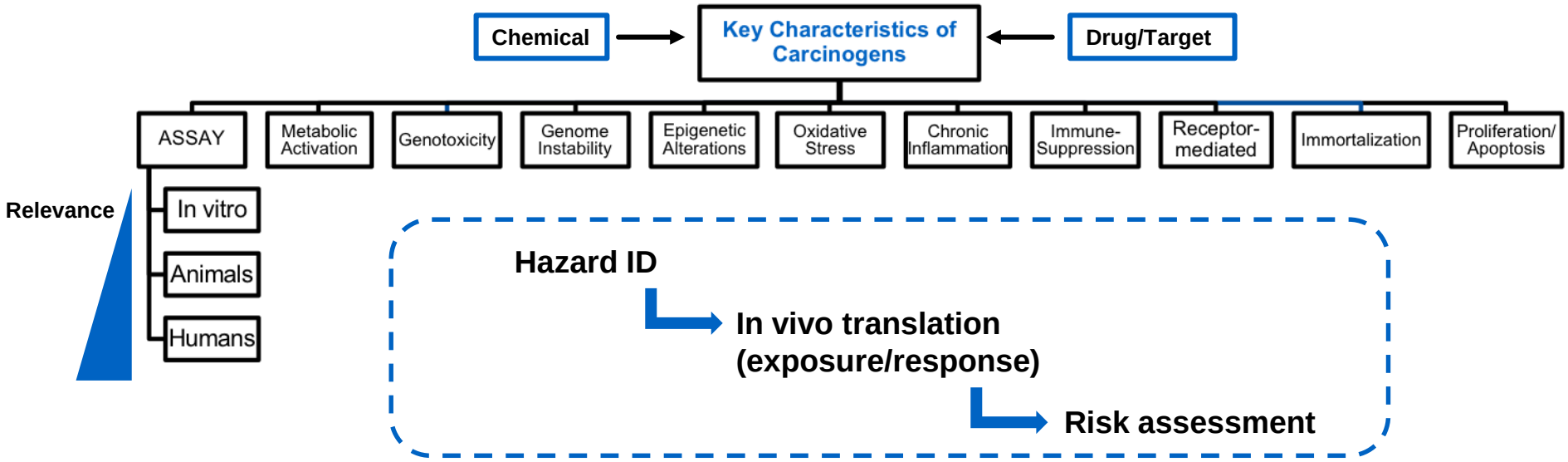
M. Fielden, Amgen

Converging Effects



M. Fielden, Amgen

Building the CarciCast Toolbox



Key Characteristics of Carcinogens can be used to define a toolbox to assess new/untested chemicals or agents for carcinogenic hazards

- Manuscript in preparation – (Fielden et al. 2019) **The Key Characteristics of Carcinogens: Relationship to the Hallmarks of Cancer and Assays and Biomarkers to Measure Them**

M. Fielden, Amgen

Challenges

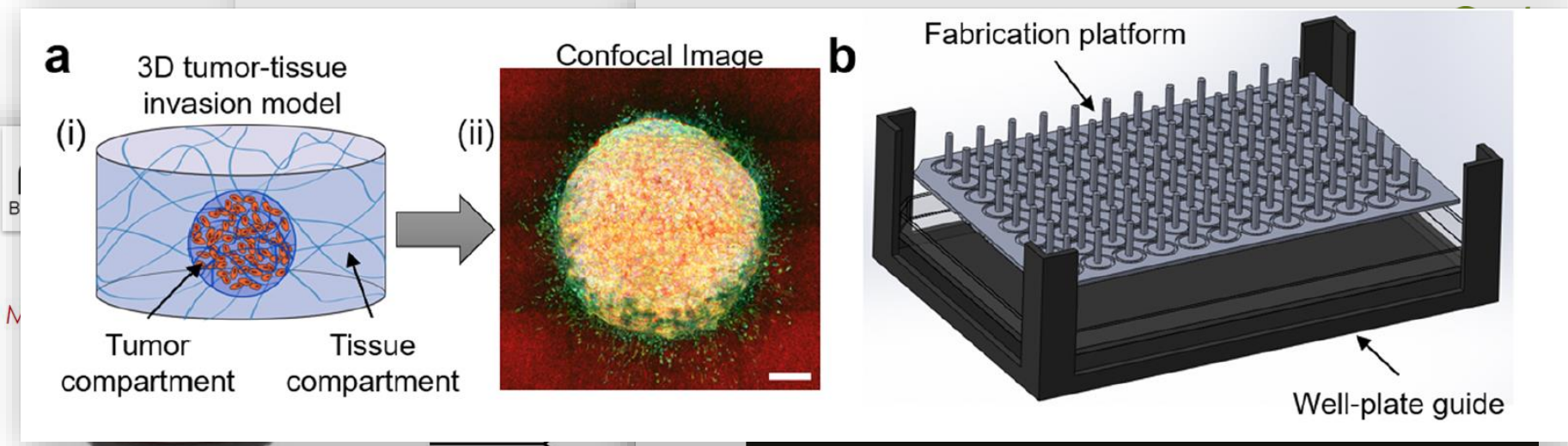
- Lack of well established or “gold standard” assays:
  - Eg. KC 6: Induces chronic inflammation, KC 5: Induces oxidative stress
- How to identify & characterize the most appropriate assays?
  - Endpoints specific and relevant to carcinogenic process?
  - Rationale for concentration/dose selection
- How to integrate results from multiple KC’s?
- How to relate in vitro results to realistic in vivo exposures?
- In vivo biomarkers of the KC’s in animals/humans needed to understand in vitro-in vivo translation and risk assessment

Measuring the Hallmarks: Complex Systems Models

- Ideal Characteristics:
  - Human-relevant platform
  - Ability to measure interdependent biological responses
  - Provide insight into tumor “tipping point”
    - Temporal and biological
  - Query impact of dose, frequency, repetition, duration, and multiplicity of exposures
  - Represent genetic differences in susceptibility, resilience, and resistance

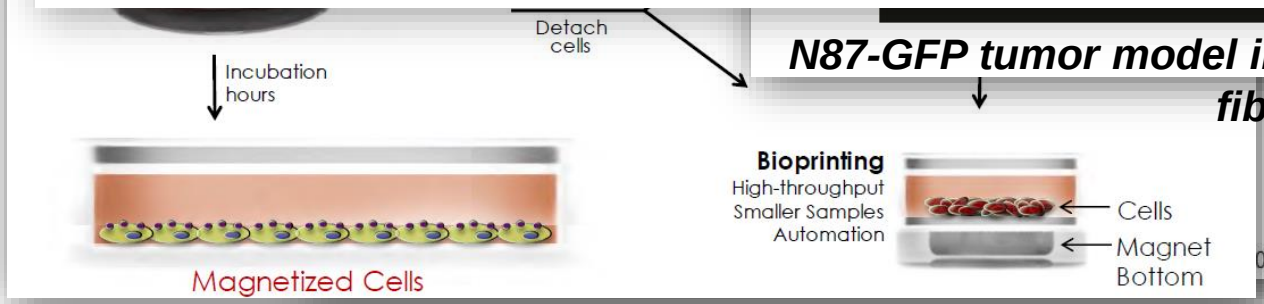
3D models: Repurposing Drug Development Platforms

3D human breast tumor models have been bioprinted
with defined multi-cellular composition and architecture



N87-GFP tumor model in c ***fibrob***

Puls et al. Nat Sci Rep 2018



BioMAP Oncology Panels

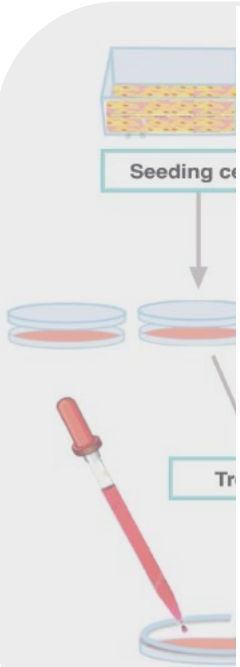
Human Primary Cells + Microenvironment

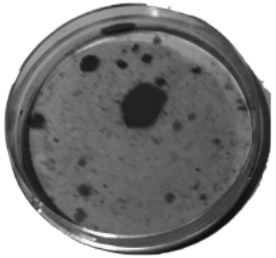
Human Biology Modeled by Oncology Panels

Panel	System		Description
Colorectal Cancer (CRC) Panel	StroHT29	Colorectal Cancer - Stro	The Colorectal Cancer - Stro (StroHT29) system models the host stromal-tumor microenvironment by capturing the complex interactions between tumor cells, the host stromal network, and infiltrating immune cells recruited into the tumor mass.
	VascHT29	Colorectal Cancer - Vasc	The Colorectal Cancer - Vasc (VascHT29) system models host vascular-tumor microenvironment by capturing the complex interactions between tumor cells, the host vascular network, and infiltrating immune cells associated with angiogenesis.
Non-Small Cell Lung Cancer (NSCLC) Panel	StroNSCLC	Lung Cancer - Stro	The Lung Cancer - Stro (StroNSCLC) host-NSCLC tumor microenvironment model system consists of human primary fibroblasts co-cultured with a NSCLC cell line, NCI-H1299, and human peripheral blood mononuclear cells. These conditions model the host stromal-tumor microenvironment by capturing the complex interactions between tumor cells, the host stromal network, and infiltrating immune cells recruited into the tumor mass.
	VascNSCLC	Lung Cancer - Vasc	The Lung Cancer - Vasc (VascNSCLC) host-NSCLC tumor microenvironment model system consists of human primary vascular endothelial cells co-cultured with a NSCLC cell line, NCI-H1299, and human peripheral blood mononuclear cells. These conditions model the host vascular-tumor microenvironment by capturing the complex interactions between tumor cells, the host vascular network, and infiltrating immune cells associated with angiogenesis.

eurofins | DiscoverX
A Eurofins Discovery Company

Transformics Assay

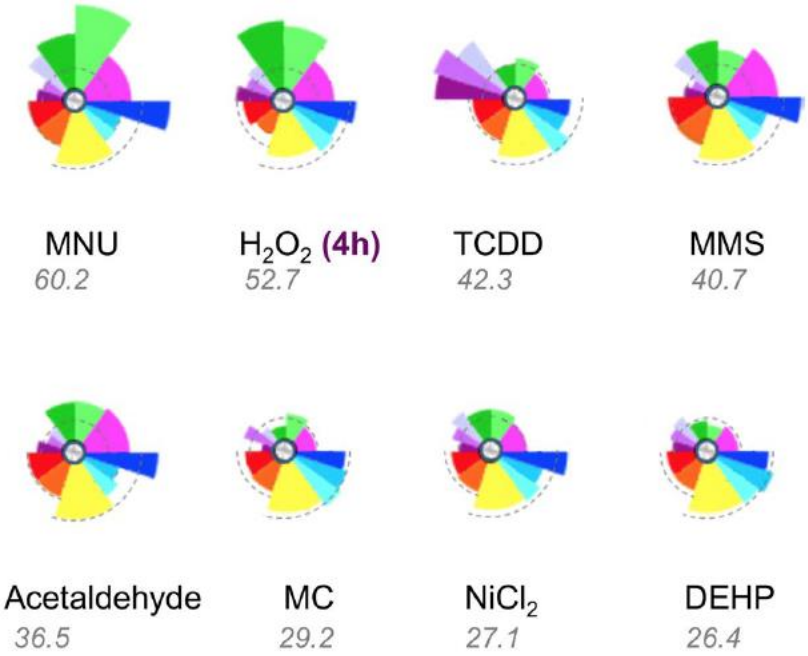


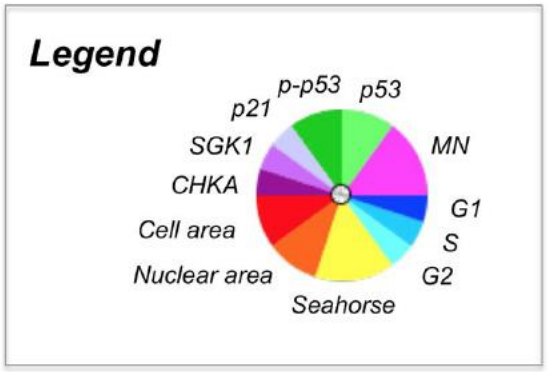
3-MCA	Main molecular endpoints			Transformation phenotypic endpoint
	24h	72h	32 days	32 days
0.04 µg/ml	• Cell adhesion mechanisms • Cell cycle regulation • Apoptosis regulation • Retinol metabolism	• Immune response • Apoptosis regulation • Estradiol metabolism	No statistically significant gene modulation	None
4 µg/ml	• **Immune response** • Cytoskeleton remodeling • Cell cycle regulation • Apoptosis regulation • Retinol metabolism	• **Immune response** • Cell cycle regulation • Estradiol metabolism	• **Immune response**	

Mascolo et al. 2018, *Carcinogenesis*

- Combining the cell transformation assay with transcriptomics
- Identify dose- and time-dependent signals, discriminate adaptive from adverse responses

Multiplexed HM-related Endpoints In Vitro

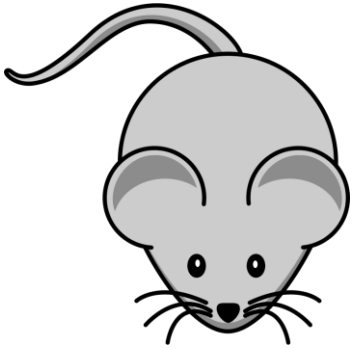




Wilde et al. Arch Tox 2018

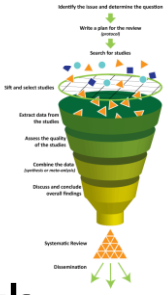
Chemical	In vivo TD_{50}	Species
TCDD	0.000023 mg/kg/day	rat
MNU	0.0927 mg/kg/day	rat
MMS	32 mg/kg/day	mouse
MC	56 mg/kg/day	rat
Acetaldehyde	153 mg/kg/day	rat
DEHP	716 mg/kg/day	rat
H2 O2	7540 mg/kg/day	mouse
NiCl2	Data unavailable	

Transgenic Mouse Models



- Ex: STAT1 deficient mouse model (STAT1KO 129S6/SvEvTac-*Stat1*^{tm1Rds}, *Borowsky lab*)
- STAT1 deficiency is a germline mutation, the emergence of tumors requires secondary mutations and/or other adaptations within the microenvironment.
- Prolonged latency supports the “adaptive oncogenesis” theory: changes in the host microenvironment facilitate the expansion of preexisting mutant populations
- Models the most common category of human breast cancer: age related (post-menopausal) ERa+ luminal carcinoma

Semi-supervised systematic review

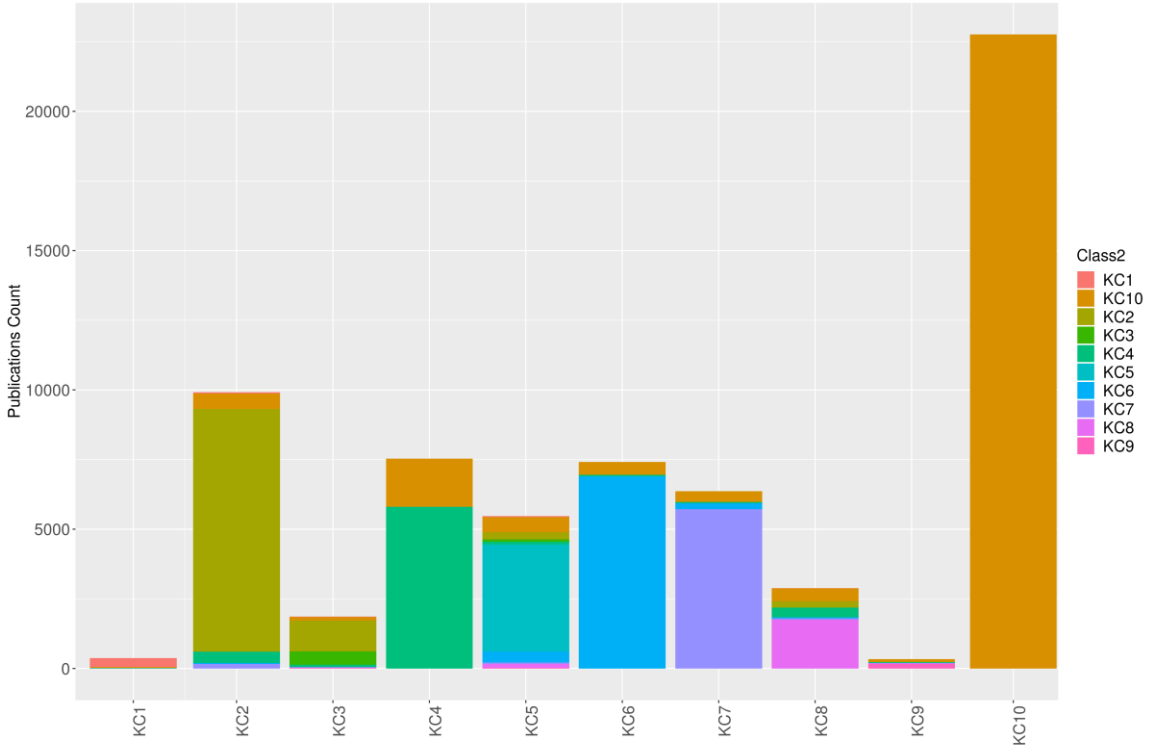


- Initial search strategy:
  Work with NTP Report on Carcinogens and Office of Health Assessment and Translation to identify keywords
  - **256** keywords mapped to HM of Cancer and KC of Carcinogens
  - **7** keywords for assays/biomarker, crossed with HM of Cancer and KC of Carcinogens
- Recruit participants to screen and tag abstracts
  - Metadata: KC, HM, Organism, Publication type, Study type
  - Mesh terms automatically tracked for PubMed articles

Initial corpus

- PubMed + Scopus database
- Literature from last 10 years

TOTAL:
- **32,605 PubMed**
- **35,171 Scopus**



PubMed publications: KC coverage

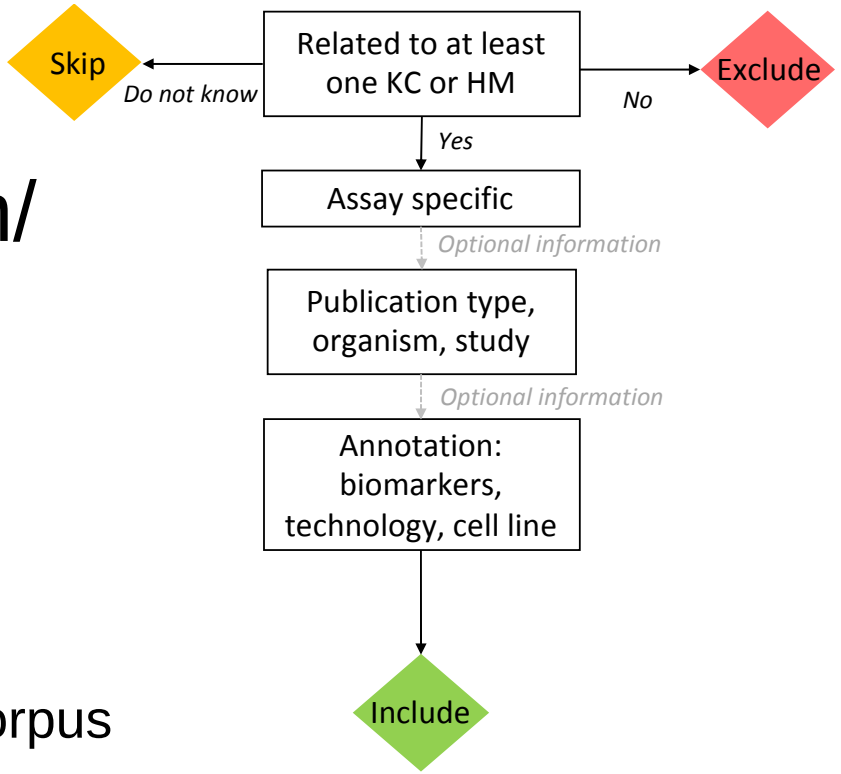
Without replicates:
57,036 publications

https://sysrev.com/

Sysrev: Semi-automated review platform

https://sysrev.com/

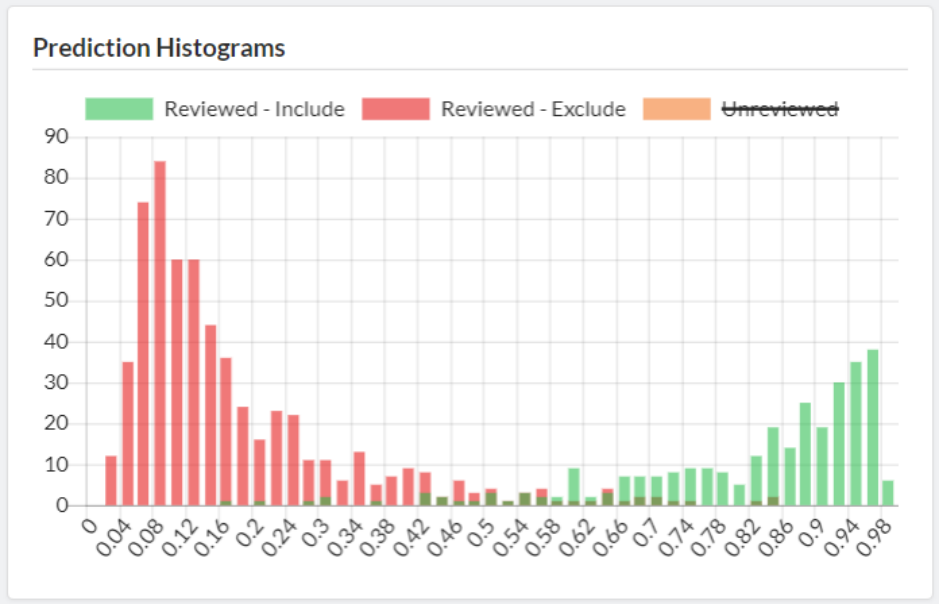
@sysrev1

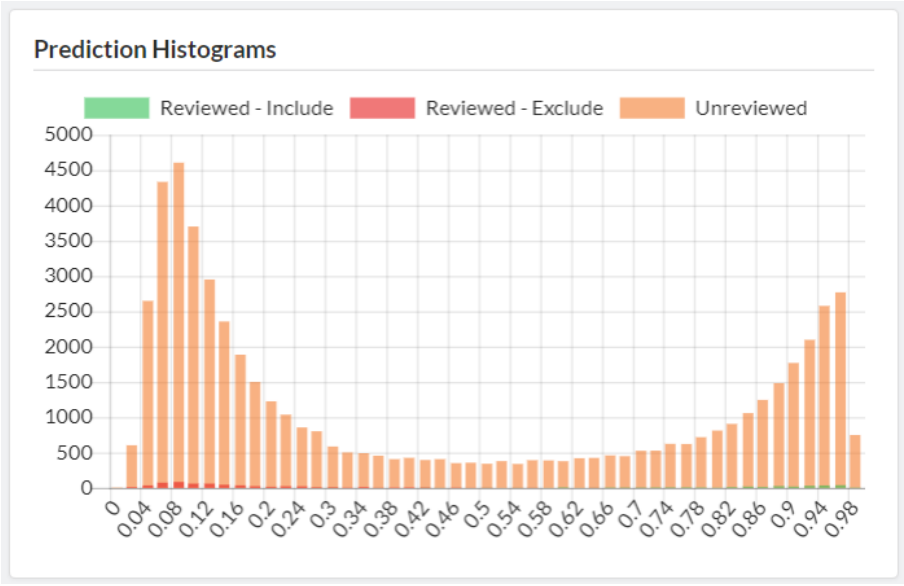


- Freely available website
- Abstract screening and annotating
- Intuitive user interface
- Including mobile/tablet access
- Uses machine learning to rank the corpus

Register	Project name	Project link
https://sysrev.com/register	Hallmark and key characteristics mapping	https://sysrev.com/p/3588

Machine Learning: Inclusion/Exclusion Models

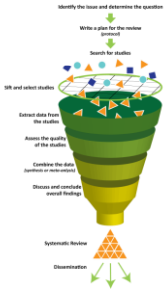




Neural network model produces a predicted score for each article
(0 not relevant, 1 relevant)

https://sysrev.com/p/3588/

Borrel et al. Poster



National Institute of Environmental Health Sciences

Systematic review: an initiative to map cancer hallmarks and key characteristics

Alexandre Borrel[1], Amy Wang[2], Lara Handler[3], N. Kleinstreuer[1,4]

[1]NIEHS/DIR/BCBB, RTP, NC; [2]NIEHS/DNTP/RoC, RTP, NC; [3]ILS, RTP, NC [4]NIEHS/DNTP/NICEATM, RTP, NC

Introduction

Carcinogenesis is a multi-step process in which normal cells are transformed into cancer cells by acquiring various properties that allow them to form tumors or malignant cancers. These acquired properties of cancer cells that distinguish them from normal cells have been classified as a series of ten Cancer Hallmarks. (Hanahan 2011). From a chemical perspective, a set of key characteristics commonly exhibited by established human carcinogens have been defined and applied by the International Agency for Research on Cancer (IARC) (Smith 2016)

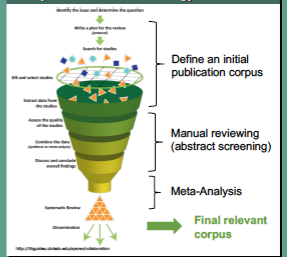
HM1: Sustained Proliferative Signaling
HM2: Evasion of Anti-growth Signaling
HM3: Resistance to Cell Death
HM4: Replicative Immortality
HM5: Angiogenesis
HM6: Tissue Invasion and Metastasis
HM7: Dysregulated Metabolism
HM8: Immune System Evasion
HM9: Genetic Instability
HM10: Inflammation

KC1: Is Electrophile or can be Activated to Electrophiles
KC2: Is Genotoxic
KC3: Activates Mutagenic DNA Repair & Promotes Genomic Instability
KC4: Induces Epigenetic Alterations
KC5: Induces Oxidative stress
KC6: Induces Chronic Inflammation
KC7: Is Immunosuppressive
KC8: Modulates Receptors-mediated effects
KC9: Causes Immortalization
KC10: Alters Cell Proliferation, Cell Death or Nutriment Supply

1. Objectives

This project is intended to support work being done by the Cancer and Environmental Mixtures Committee: Assay and Biomarker Subgroup, as well as support the NTP's Strategic Health Effects Innovation on Carcinogenicity Testing for the 21st Century. The aim of this literature review is to identify novel assays and biomarkers that map to the hallmarks of cancer (HM) and the key characteristics of carcinogens (KC). The overarching goals of developing such a literature database include informing new testing strategies and frameworks to incorporate mechanistic data into cancer risk assessment and developing effective screening tools to detect the carcinogenic potential of environmental chemicals (including mixtures). Other downstream applications could ultimately include engineering safer products and designing more effective multi-target therapeutics.

2. Systematic review strategy

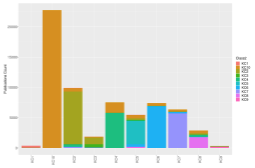


3. Initial corpus

With the NTP Report on Carcinogens and Office of Health Assessment and Translation, we identified **256 keywords mapped to HM of Cancer and KC of Carcinogens** crossed with **7 keywords for assays/biomarkers**. To identify the most relevant cutting-edge technologies (or those that are still in widespread use), only publications after 2008 were included, and book chapters/dissertation/thesis were excluded. Keyword publication searches were performed on both PubMed and Scopus database.

	PubMed	Scopus
Count of publications	32,605	35,171

Numerous publications were found in both databases. Distribution of KC keywords across publications is shown in the histogram below.



Following the literature search, the initial corpus includes **~57,000 publications** from Scopus and PubMed

4. Sysrev platform

sysrev — **https://sysrev.com**

Semi-automated review platform
- **Freely** available website
- **Abstract screening** and annotating
- **Intuitive** user interface, including mobile/tablet access
- Uses **machine learning** to rank the corpus

5. Abstract screening

Sysrev provides a web interface to label abstracts. Seven labels can be specified for each abstract as well as biomarker and technology annotation. Each article need two positive reviews to be included in the corpus. In case of conflict a third reviewer need to resolve it.

- Include
- Assay specified (technology specified)
- Hallmarks
- Key characteristics
- Organisms
- Publication type
- Type of study

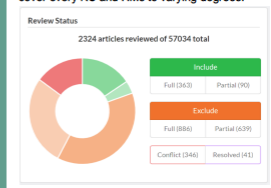
Assay technology can be annotated in the abstract as well as biomarker

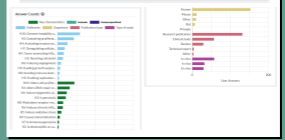
6. Machine learning

Using a neural network approach, for each publication Sysrev computes an inclusion score based on keywords in publications previously included in the corpus. The model is continuously iterating and improving based on reviewer feedback. See below for a distribution for reviewed and unreviewed publications.

7. Preliminary results

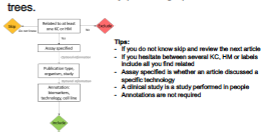
As of early April 2019, 2324 publications have been reviewed. Only 15% of the publication reviewed are included in the final corpus, and cover every KC and HMs to varying degrees.





8. Support

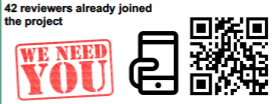
We provide reviewer support for abstract screening with a Standard Operating Procedure available on the platform. Periodic webinars help standardize reviews by providing tips and decision trees.



Tips:
- If you do not know skip and review the next article
- If you hesitate between several KC, HM or labels include all you find related
- Assay specified is whether an article discussed a specific technology
- A clinical study is a study performed in people
- Annotations are not required

Conclusion

42 reviewers already joined the project

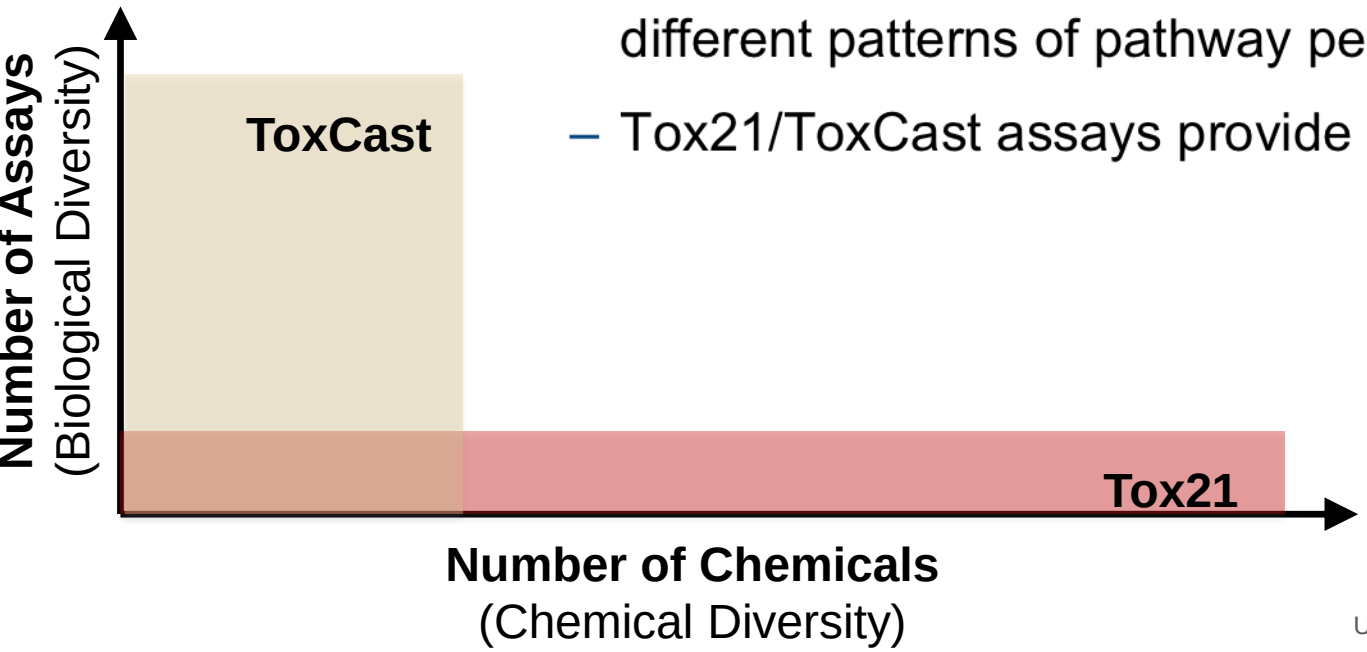


Join us now!

National Institutes of Health • U.S. Department of Health and Human Services

Applying HTS data

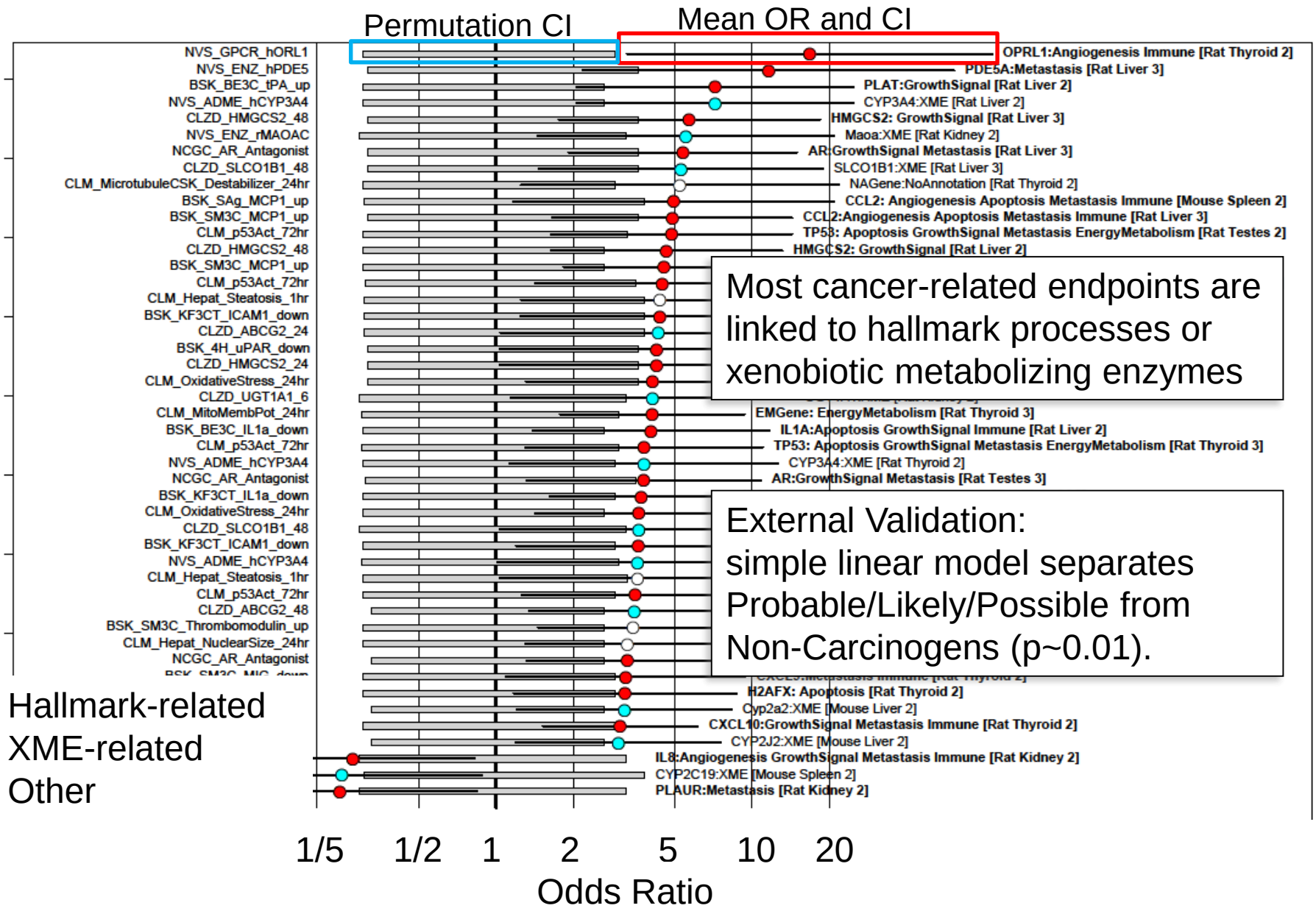
- We observe that some chemicals perturb multiple cancer hallmark pathways
- Hypothesis: A chemical that perturbs many pathways related to cancer hallmark processes will be more likely to cause cancer in the lifetime of an animal than a chemical that perturbs few such pathways
  - Chemicals can increase cancer risk through many different patterns of pathway perturbations
  - Tox21/ToxCast assays provide decent HM coverage



Initial Approach

- Link Tox21/Toxcast assays with genes, pathways, cancer hallmarks
  - Use published pathways and Gene Ontology keywords
- Calculate univariate associations
  - *In vitro* assay x *in vivo* cancer endpoints, odds-ratio (OR)
  - Multiple testing corrections with permutation tests
  - Keep associations with OR>2, Lower Confidence Interval > 1
- Rank chemicals by number of hits
- Forward validate with 60 chemicals not in signature development set

Kleinstreuer et al. 2013, Tox Sci



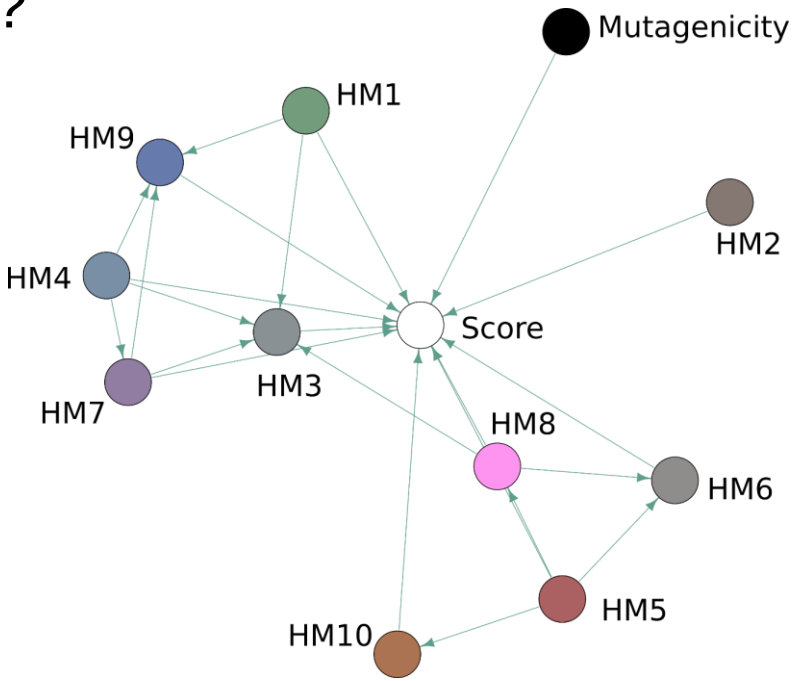
Most cancer-related endpoints are linked to hallmark processes or xenobiotic metabolizing enzymes

External Validation:
simple linear model separates Probable/Likely/Possible from Non-Carcinogens (p~0.01).

- Hallmark-related
- XME-related
- Other

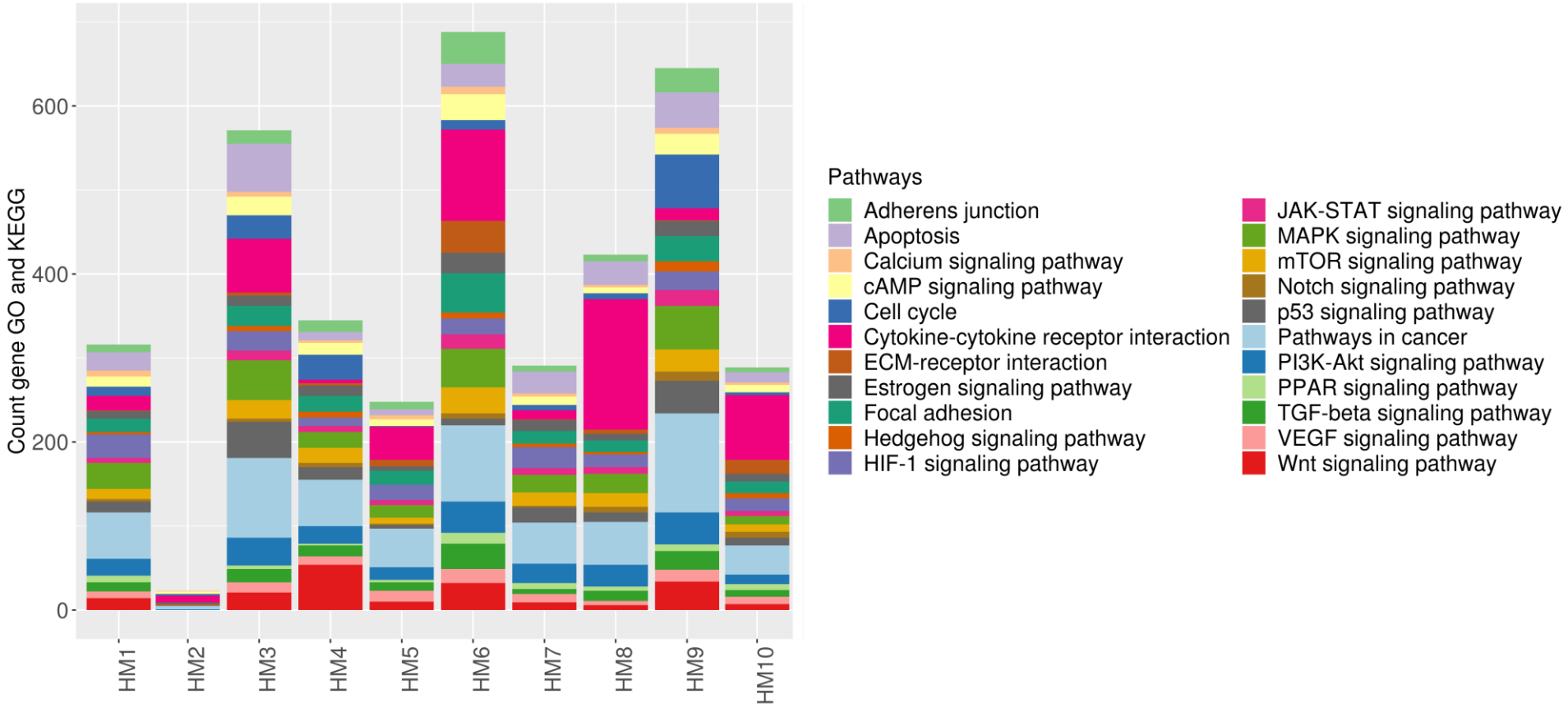
Current Approach: Biologically-based Bayesian Networks

- Bayesian Networks provide a probabilistic means to predict an outcome based on measured values
- Can this approach be used to predict chemical carcinogenicity potential from hallmark-related *in vitro* and *in silico* assays?



Hallmark/Gene/Assay mapping

- Identify updated (2019) ToxCast/Tox21 assay targets (~350) mapped to hallmark-related genes



Borrel et al. Poster

Systematic review: an initiative to map cancer hallmarks and key characteristics

Alexandre Borrel[1], Amy Wang[2], Lara Handler[3], N. Kleinstreuer[1,4]
[1]NIEHS/DIR/BCBB, RTP, NC; [2]NIEHS/DNTP/RoC, RTP, NC; [3]ILS, RTP, NC [4]NIEHS/DNTP/NICEATM, RTP, NC

Introduction

Carcinogenesis is a multi-step process in which normal cells are transformed into cancer cells by acquiring various properties that allow them to form tumors or malignant cancers. These acquired properties of cancer cells that distinguish them from normal cells have been classified as a series of ten Cancer Hallmarks. (Hanahan 2011). From a chemical perspective, a set of key characteristics commonly exhibited by established human carcinogens have been defined and applied by the International Agency for Research on Cancer (IARC) (Smith 2016)

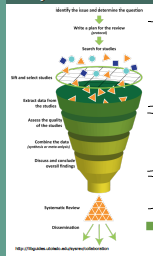
HM1: Sustained Proliferative Signaling
HM2: Evasion of Anti-growth Signaling
HM3: Resistance to Cell Death
HM4: Replicative Immortality
HM5: Angiogenesis
HM6: Tissue Invasion and Metastasis
HM7: Dysregulated Metabolism
HM8: Immune System Evasion
HM9: Genetic Instability
HM10: Inflammation

KC1: Is Electrophile or can be Activated to Electrophiles
KC2: Is Genotoxic
KC3: Activates Mutagenic DNA Repair & Promotes Genomic Instability
KC4: Induces Epigenetic Alterations
KC5: Induces Oxidative stress
KC6: Induces Chronic Inflammation
KC7. Is Immunosuppressive
KC8: Modulates Receptors-mediated effects
KC9: Causes Immortalization
KC10: Alters Cell Proliferation, Cell Death or Nutriment Supply

1. Objectives

This project is intended to support work being done by the Cancer and Environmental Mixtures Committee: Assay and Biomarker Subgroup, as well as support the NTP's Strategic Health Effects Innovation on Carcinogenicity Testing for the 21st Century. The aim of this literature review is to identify novel assays and biomarkers that map to the hallmarks of cancer (HM) and the key characteristics of carcinogens (KC). The overarching goals of developing such a literature database include informing new testing strategies and frameworks to incorporate mechanistic data into cancer risk assessment and developing effective screening tools to detect the carcinogenic potential of environmental chemicals (including mixtures). Other downstream applications could ultimately include engineering safer products and designing more effective multi-target therapeutics.

2. Systematic review strategy

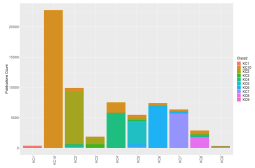


3. Initial corpus

With the NTP Report on Carcinogens and Office of Health Assessment and Translation, we identified **256 keywords mapped to HM of Cancer and KC of Carcinogens** crossed with **7 keywords for assays/biomarkers**. To identify the most relevant cutting-edge technologies (or those that are still in widespread use), only publications after 2008 were included, and book chapters/dissertation/thesis were excluded. Keyword publication searches were performed on both PubMed and Scopus database.

	PubMed	Scopus
Count of publications	32,605	35,171

Numerous publications were found in both databases. Distribution of KC keywords across publications is shown in the histogram below.



Following the literature search, the initial corpus includes **~57,000 publications** from Scopus and PubMed

4. Sysrev platform

https://sysrev.com

Semi-automated review platform
- **Freely** available website
- **Abstract screening** and annotating
- **Intuitive** user interface, including mobile/tablet access
- Uses **machine learning** to rank the corpus

5. Abstract screening

Sysrev provides a web interface to label abstracts. Seven labels can be specified for each abstract as well as biomarker and technology annotation. Each article need two positive reviews to be included in the corpus. In case of conflict a third reviewer need to resolve it.

1 Include
2 Assay specified (technology specified)
3 Hallmarks
4 Key characteristics
5 Organisms
6 Publication type
7 Type of study

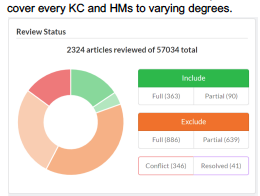
Assay technology can be annotated in the abstract as well as biomarker

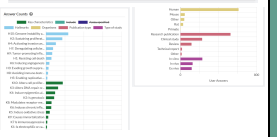
6. Machine learning

Using a neural network approach, for each publication Sysrev computes an inclusion score based on keywords in publications previously included in the corpus. The model is continuously iterating and improving based on reviewer feedback. See below for a distribution for reviewed and unreviewed publications.

7. Preliminary results

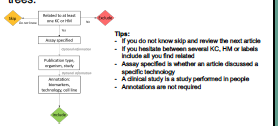
As of early April 2019, 2324 publications have been reviewed. Only 15% of the publication reviewed are included in the final corpus, and cover every KC and HMs to varying degrees.





8. Support

We provide reviewer support for abstract screening with a Standard Operating Procedure available on the platform. Periodic webinars help standardize reviews by providing tips and decision trees.



Conclusion

42 reviewers already joined the project

Join us now!

National Institutes of Health • U.S. Department of Health and Human Services

QSAR models for Key Characteristics

A Section 508–conformant HTML version of this article is available at http://dx.doi.org/10.1289/ehp.1509912.

Review

Key Characteristics of Carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis

Martyn T. Smith,[1] Kathryn Z. Guyton,[2] Catherine F. Gibbons,[3] Jason M. Fritz,[3] Christopher J. Portier,[4] Ivan Rusyn,[5] David M. DeMarini,[3] Jane C. Caldwell,[3] Robert J. Kavlock,[3] Paul F. Lambert,[6] Stephen S. Hecht,[7] John R. Bucher,[8] Bernard W. Stewart,[9] Robert A. Baan,[2] Vincent J. Cogliano,[3] and Kurt Straif[2]*

[1]Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, California, USA; [2]International Agency for Research on Cancer, Agency, Washington, DC, USA, and Resea [5]Department of Veterinary Integrative Bios College Station, Texas, USA; [6]McArdle Lab Madison, Wisconsin, USA; [7]Masonic Cancer Program, National Institute of Environmen Research Triangle Park, North Carolina, U

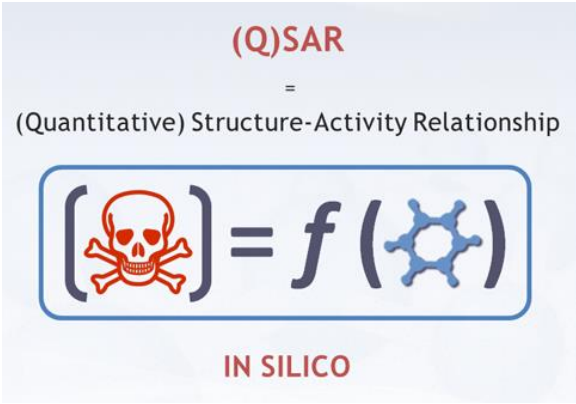
ORIGINAL ARTICLE

Application of the key characteristics of carcinogens in cancer hazard identification

Kathryn Z.Guyton[1], Ivan Rusyn[2], Weihsueh A.Chiu[2], Denis E.Corpet[3], Martin van den Berg[4], Matthew K.Ross[5], David C.Christiani[6,7], Frederick A.Beland[8] and Martyn T.Smith[9,*]

Use of high-throughput in vitro toxicity screening data in cancer hazard evaluations by IARC Monograph Working Groups

Weihsueh A. Chiu[1,*], Kathryn Z. Guyton[2], Matthew T. Martin[3], David M. Reif[4], and Ivan Rusyn[1,*]

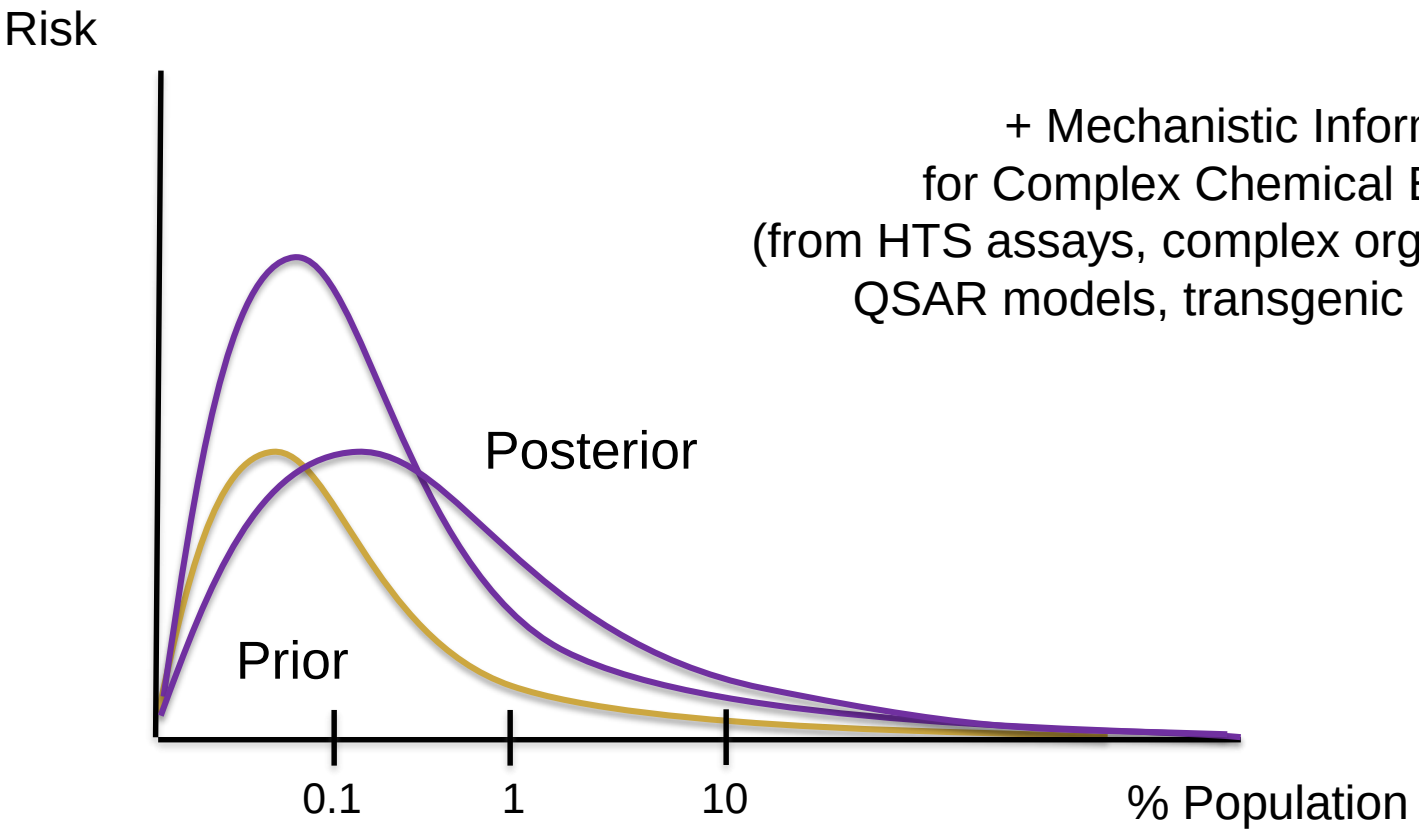


- Use KC mapping of HTS assays to identify training set chemicals (active/inactive) for each KC-QSAR model

Ongoing Work

- Include ToxCast assays without specific gene targets (e.g. proliferation, mitochondrial function)
- Refine scoring metrics, investigate tissue-specific endpoints, feature selection algorithms to id minimum assay set, id targets missing from HTS, investigate mis-predicted chemicals
- Incorporate informative priors based on systematic literature review results into BN learning
- Combine with KC-QSARs and low-throughput complex mechanistic assays to form integrated testing strategies
- Ultimate goal: probabilistic chemical (complex mixture) screening for carcinogenicity using battery of *in vitro* and *in silico* tests.

Addressing Carcinogenic Risk Probabilistically



QUESTIONS?